

Draft Genome Sequence of *Lactobacillus oryzae* Strain SG293^T

Yasuhiro Tanizawa,^{a,b} Takatomo Fujisawa,^b Takako Mochizuki,^b Eli Kaminuma,^b Yasukazu Nakamura,^b Masanori Tohno^c

Department of Computational Biology, Graduate School of Frontier Sciences, the University of Tokyo, Chiba, Japan^a; Genome Informatics Laboratory, National Institute of Genetics, Mishima, Japan^b; National Agriculture and Food Research Organization, National Institute of Livestock and Grassland Science, Nasushiobara, Japan^c

We report the 1.86-Mb draft genome and annotation of *Lactobacillus oryzae* SG293^T isolated from fermented rice grains. This genome information may provide further insights into the mechanisms underlying the fermentation of rice grains.

Received 1 August 2014 Accepted 4 August 2014 Published 28 August 2014

Citation Tanizawa Y, Fujisawa T, Mochizuki T, Kaminuma E, Nakamura Y, Tohno M. 2014. Draft genome sequence of *Lactobacillus oryzae* strain SG293^T. *Genome Announc.* 2(4): e00861-14. doi:10.1128/genomeA.00861-14.

Copyright © 2014 Tanizawa et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Masanori Tohno, tohno@affrc.go.jp.

In the livestock industry, lactic acid bacteria (LAB) are widely used as inoculants for silage fermentation. Recently, for livestock feed, rice grains have gained attention as an alternative to corn (1). During an investigation of LAB from fermented rice grains, *Lactobacillus oryzae*, a novel species of the genus *Lactobacillus*, was isolated (2). It may be associated with adequate fermentation of rice grain silage and has the potential for biotechnological application, such as an effective inoculant for rice grains. Here we report the draft genome sequence of strain SG293, which is a type strain of *L. oryzae*.

The genomic DNA of SG293^T was extracted and purified from a culture grown in De Man, Rogosa, and Sharpe broth (Difco) using QIAGEN Genomic tips (Qiagen). The whole genome was sequenced using the 250-bp pair-end Illumina MiSeq system (Fugen, Japan), yielding a total of 8,545,944 reads, which corresponds to approximately 1,100-fold coverage. Reads that contained adaptor sequences were filtered using the cutadapt software (3). *De novo* assembly was performed using Platanus version 1.2.1 (4). Contigs shorter than 300 bp were eliminated.

The draft genome was annotated by the Microbial Genome Annotation Pipeline (MiGAP) online annotation server (5), which combines prediction of open reading frames by MetaGeneAnnotator 1.0 (6), tRNA prediction by tRNAscan-SE 1.23, rRNA prediction by RNAmmer 1.2, and homology searches against the RefSeq, TrEMBL, and Clusters of Orthologous Groups (COGs) protein databases. The KEGG Automatic Annotation Server (KAAS) (7) was also used to search for metabolic pathways. In addition to these automated annotations, protein sequences were queried against the NCBI-nr and Swiss-Prot databases using BLASTp, searched against signature/domain databases using InterProScan, and part of the annotation was curated manually. Clustered regularly interspaced short palindromic repeats (CRISPR) loci were searched using the CRISPRfinder server (8). The annotated genome was submitted to the GenomeRefine web service (<http://genome.annotation.jp/genomerefine/>), which assists with the refinement of the annotation and registration at the International Nucleotide Sequence Database Collaboration (INSDC) through DDBJ.

The draft genome sequence of *L. oryzae* SG293^T consists of 93 contigs and includes 1,860,394 bp with an overall G+C

content of 43%. In the genome, 1,891 protein-coding sequences (CDSs) as well as 59 copies of tRNA and six copies of rRNA were predicted. Approximately 73% of the CDSs were assigned to specific COGs and 53% were assigned a KEGG orthology (KO) number. All enzymes required for the phosphoketolase pathway are encoded in the genome, whereas some genes involved in the Embden–Meyerhof–Parnas pathway are absent, supporting the classification of this strain as an obligate heterofermentative LAB. One CRISPR locus was detected in the genome, which contained three CRISPR-associated proteins (LOSG293_470040–470060) and 47 spacer sequences separated by 36-bp direct repeats. This genomic information may provide further insights into the mechanisms underlying the fermentation of rice grains.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession no. [BBJM00000000](http://www.ncbi.nlm.nih.gov/nuccore/BBJM00000000). The version described in this paper is the first version, BBJM01000000.

ACKNOWLEDGMENTS

We thank M. Ezure and M. Arita for technical support and helpful discussions. The computation in this work was done using the super computer system of the National Institute of Genetics, the Research Organization of Information and Systems.

This study was supported in part by a Grant-in-Aid for Young Scientists (A) (25712032) from the Japan Society for the Promotion of Sciences and an NIG Collaborative Research Program (2012–2088, 2013–2070) from the National Institute of Genetics.

REFERENCES

1. Miyaji M, Matsuyama H, Hosoda K. 2014. Effect of substituting brown rice for corn on lactation and digestion in dairy cows fed diets with a high proportion of grain. *J. Dairy Sci.* 97:952–960. <http://dx.doi.org/10.3168/jds.2013-7046>.
2. Tohno M, Kitahara M, Irisawa T, Inoue H, Uegaki R, Ohkuma M, Tajima K. 2013. *Lactobacillus oryzae* sp. nov., isolated from fermented rice grain (*Oryza sativa* L. subsp. *japonica*). *Int. J. Syst. Evol. Microbiol.* 63: 2957–2962. <http://dx.doi.org/10.1099/ijsm.0.048918-0>.
3. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12. <http://dx.doi.org/10.14806/ej.17.1.200>.
4. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A.

- Hayashi T, Itoh T. 2014. Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* 24: 1384–1395. <http://dx.doi.org/10.1101/gr.170720.113>.
5. Sugawara H, Ohyama A, Mori H, Kurokawa K. 2009. Microbial genome annotation pipeline (MiGAP) for diverse users, poster S001-1-2. 20th Int. Conf. Genome Informatics (GIW2009), Yokohama, Japan.
6. Noguchi H, Taniguchi T, Itoh T. 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res.* 15:387–396. <http://dx.doi.org/10.1093/dnares/dsn027>.
7. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35:W182–W185. <http://dx.doi.org/10.1093/nar/gkm321>.
8. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35:W52–W57. <http://dx.doi.org/10.1093/nar/gkm360>.